Chapitre 3 : Théorèmes limites

Introduction : rappels sur la valeur absolue

♡ **PARAGRAPHE IMPORTANT !!!!!!!!**

$|x-y|$ représente la distance $d(x,y)$ entre deux points de l'axe réel.

La distance entre deux points est nulle ssi $x=y$. C'est une quantité symétrique : $d(y,x)=d(x,y)$. Soit r un réel positif ou nul (**le rayon**). L'ensemble $\{t \quad \text{tq} \quad |t-m| \leq r\}$ est l'intervalle I de **centre** m et de **rayon** r : c'est l'ensemble des points t dont la distance au centre m est inférieure ou égale à r.
On trouve facilement que les bornes sont $m \pm r$, c'est à dire $I=[m-r, m+r]$.

I- La loi des grands nombres et l'inégalité de Bienaymé-Tchebychev

1. Loi (forte) des grands nombres (Énoncé admis)

Ici loi signifie théorème.

Enoncé. *Soit une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes de même loi. On suppose que $\mathbb{E}(X_1)$ existe. On pose comme d'habitude $S_n=\sum_{k=1}^{n} X_k$. Alors $\dfrac{S_n}{n}$ converge presque sûrement (*) vers $\mathbb{E}(X_1)$.*

Exemple : on lance n fois un dé non truqué. On compte le nombre de fois que le six apparaît. Si on définit $X_i=\mathbb{1}_{\{\text{le premier lancer donne } 6\}}$, S_n est justement le nombre de six et d'après la loi des grands nombres, la fréquence $\frac{S_n}{n}$ des six tend vers $\mathbb{E}(X_1)=\mathbb{P}(\text{le i-ème lancer donne } 6)=\frac{1}{6}$.
La probabilité d'un événement peut donc s'interpréter comme une limite de fréquences.

2. Loi faible des grands nombres

Comme son nom l'indique, c'est une version plus simple de la loi des grands nombres. Il est même possible de donner en L2 une démonstration complète de ce théorème.

♡ **Enoncé.** *Soit une suite $(X_k)_{k\geq 1}$ de variables aléatoires indépendantes de même loi. On suppose que $\mathbb{E}(X_1)$ et $\mathrm{var}(X_1)$ existent. On pose comme d'habitude $S_n=\sum_{k=1}^{n} X_k$. Alors pour tout réel α strictement positif (et qui ne dépend pas du hasard) on a l'inégalité (dite* **inégalité de Bienaymé-Tchebychev** *)*

$$\mathbb{P}(|S_n-\mathbb{E}(S_n)|>\alpha) \leq \frac{\mathrm{var}(S_n)}{\alpha^2}.$$

♡ **Exemple fondamental**
On suppose que S_n suit la loi binômiale $\mathcal{B}(n,p)$. Alors $\mathbb{P}\left(|S_n-np|>\alpha\right) \leq \dfrac{npq}{\alpha^2}$.

On en déduit en posant $\beta=\frac{\alpha}{n}$ le

(*) *Explications complémentaires en amphi.*

Corollaire 1. $\mathbb{P}\left(\left| \frac{S_n}{n} - p \right| > \beta \right) \leq \frac{pq}{n\beta^2}$.

Conséquence de ce corollaire.

Quand n tend vers l'infini, le majorant $\frac{pq}{n\beta^2}$ tend vers 0. Cela prouve que $\frac{S_n}{n}$ tend à se concentrer en p : plus précisément, si r est un rayon fixé, la probabilité que $\frac{S_n}{n}$ appartienne à l'intervalle de centre p et de rayon r est de plus en plus proche de 1 quand n est grand.

Cette inégalité sert à majorer la probabilité d'événements anormaux $\{\frac{S_n}{n}$ tend à s'écarter de $p\}$.

Complément. Variables aléatoires de carré intégrable.

Définition. On dit dit qu'une va. réelle est de **carré intégrable** ssi $\mathbb{E}(X^2)$ est finie. Cela équivaut à dire qu'elle possède une variance finie. Ceci se prouve grâce à l'inégalité de Cauchy-Schwarz :

Inégalité de Cauchy-Schwarz. $\mathbb{E}(X)^2 \leq \mathbb{E}(X^2)$.

II- Le théorème central limite (admis)

1. Enoncé

♡ **Théorème central limite (en abrégé TCL).** *Soit une suite $(X_k)_{k\geq 1}$ de variables aléatoires indépendantes de même loi. On suppose que $\mathbb{E}(X_1)$ et var(X_1) existent et que var$(X_1) > 0$ (c'est à dire X_1 de carré intégrable. On pose comme d'habitude $S_n = \sum_{k=1}^{n} X_k$. Alors pour tous les réels $a \leq b$, quand n tend vers l'infini,*

$$\mathbb{P}\left(a \leq \frac{S_n - n\mathbb{E}(X_1)}{\sqrt{n \text{ var}(X_1)}} \leq b \right) \quad \textit{converge vers} \quad \int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})\, dt.$$

On verra dans le chapitre suivant que la fonction $t \longrightarrow \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})$ est la densité de la variable aléatoire gaussienne centrée réduite. La courbe représentative de la fonction $\frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$ s'appelle la **courbe en cloche de Gauss**. (voir le dessin page 31)

Remarque : se souvenir que les variables $Z_n = \frac{S_n - n\mathbb{E}(X_1)}{\sqrt{n\ \mathrm{var}(X_1)}}$ figurant dans le TCL sont centrées et réduites puisque $\mathbb{E}(S_n) = n\,\mathbb{E}(X_1)$ et $\mathrm{var}(S_n) = n\,\mathrm{var}(X_1)$.

2. Cas particulier (Moivre-Laplace)

On suppose que la loi de X_1 est de Bernoulli $\mathcal{B}(1,p)$ avec $0 < p < 1$. Ici S_n a la loi $\mathcal{B}(n,p)$. Le théorème central limite prend donc la forme

$$\mathbb{P}(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b) \ \text{ converge vers } \ \int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})\, dt.$$

3. Utilisation pratique du théorème central limite: l'approximation gaussienne

♡ Supposons que S_n a pour loi $\mathcal{B}(n,p)$ et $a < b$. Quand n est grand, on approxime $\mathbb{P}(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b)$ par sa limite dans le TCL $\displaystyle\int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})\, dt$.

Critère pratique : si $np > 10$ et $nq > 10$, l'approximation est suffisante pour les besoins usuels???
Ce critère exprime que plus p s'éloigne de $\frac{1}{2}$, plus n doit être grand. C'est facile à expliquer : quand p est loin de $\frac{1}{2}$, les p_k sont fortement asymétriques, ce qui donne une courbe bien différente de la courbe en cloche.

♡ **Avertissement, ne pas perdre de temps en tentant des calculs inutiles : On ne connaît pas de primitive explicite de la fonction** $t \longrightarrow \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})$!!!
On se sert donc d'une calculatrice ou d'une table.

4. Utilisation de la table distribuée en L2.

Définition. On pose pour tout x, $\Psi(x) = \displaystyle\int_0^x \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})\, dt$. Donc la limite dans l'énoncé du TCL $\displaystyle\int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})\, dt$ est $\Psi(b) - \Psi(a)$ (formule de Chasles).

On lit dans la table des valeurs approchées de $\Psi(x)$ pour $x \geq 0$. Cela représente une valeur approchée de l'aire sous la courbe en cloche entre 0 et x (hachurée ci-dessus).

Mode d'emploi : à l'intersection de la ligne 2,1 et de la colonne 7 (= 0.07), on lit la valeur approchée de $\Psi(2,17) \approx 0,4850$.

5. Étude de la fonction Ψ.

Grâce au théorème de dérivation d'une intégrale fonction de sa borne supérieure, on trouve que Ψ est dérivable avec $\Psi'(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$.
Donc Ψ est strictement croissante, impaire.
On admet que

$$\lim_{x \to +\infty} \Psi(x) = \frac{1}{2}.$$

Cela équivaut à la formule expliquée dans le chapitre suivant :

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})\, dt = 1.$$

Tableau de variation et graphe de Ψ :

6. Compléments sur la lecture de table

Si $x \geq 4$, on prend $\Psi(x) \approx 0.5000$. De même par imparité, si $x \leq -4$, on prend $\Psi(x) \approx -0.5000$.
Pour les valeurs négatives, on utilise l'imparité de Ψ.

7. Lois gaussiennes (première approche)

♡ **Définition.** On dit qu'une variable aléatoire Z possède la loi normale centée réduite ssi pour tous les réels $a \leq b$, $\mathbb{P}(a \leq Z \leq b) = \displaystyle\int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})\, dt$.

On dit aussi loi « gaussienne centrée réduite », en abrégé $\mathcal{N}(0,1)$.
Autrement dit, $\mathbb{P}(a \leq Z \leq b) = \Psi(b) - \Psi(a)$.
La table s'appelle donc un peu abusivement table de la loi $\mathcal{N}(0,1)$.
♡ **Définition.** Soient m un réel quelconque et σ (*sigma*) un réel strictement positif. On dit qu'une variable aléatoire X possède la loi normale de moyenne m et d'écart-type σ ssi $Z = \frac{X-m}{\sigma}$ est une variable aléatoire de loi $\mathcal{N}(0,1)$.

Remarques : a) On verra dans le chapitre suivant que m est l'espérance de X et que σ est bien son écart-type. Par conséquent, Z est la variable aléatoire obtenue en centrant et en réduisant X.
b) Les calculs pratiques se font en revenant à la table $\mathcal{N}(0,1)$.
Si $u \leq v$, $u \leq X \leq v$ équivaut à $\frac{u-m}{\sigma} \leq Z \leq \frac{v-m}{\sigma}$.

Donc $\mathbb{P}(u \leq X \leq v) = \mathbb{P}(\frac{u-m}{\sigma} \leq Z \leq \frac{v-m}{\sigma}) = \Psi(\frac{v-m}{\sigma}) - \Psi(\frac{u-m}{\sigma})$.

III- Recherche d'intervalles de confiance pour une proportion

Il y a deux méthodes. En général, le choix est imposé par l'énoncé.

1. Position du problème

On désire connaître la valeur p inconnue d'une probabilité d'un événement, par exemple la probabilité qu'un dé suspect donne le six. En faisant un nombre limité d'expériences, il n'est pas possible de trouver la valeur exacte de p.
On va trouver à la place une « fourchette » pour p, c'est à dire un intervalle I. Pour cela on utilise n données expérimentales. Comme elles peuvent être aberrantes (jour de malchance par exemple), il faut bien voir que I ne contient pas forcément p. On n'aura qu'une certitude probabiliste, c'est à dire que I contient p avec une probabilité assez grande.

Définition. On appelle **intervalle de confiance** au niveau de confiance $1-\alpha$ un intervalle I qui contient p avec une probabilité plus grande que le $1-\alpha$. Les bornes de l'intervalles doivent être des variables aléatoires qui ne dépendent que des variables aléatoires observées, de n et du niveau $1-\alpha$ (surtout pas de p qui est inconnu et le restera).
La définition plus correcte demande que pour tout p dans $]0,1[$, $\mathbb{P}(p \in I) \geq 1-\alpha$.
En fait, p figure dans la loi de probabiilité, puisque les calculs sont effectués en supposant S_n de loi $\mathcal{B}(n,p)$.

On utilise le plus souvent les niveaux $1-\alpha$ égaux à 0.90, 0.95 et 0.99.
Remarque : on note de façon compliquée $1-\alpha$ le niveau de confiance par référence à la quantité (petite) α qui apparaît dans la théorie des tests d'hypothèses.

2. Hypothèses de travail

Soit p une probabilité inconnue. Soit S_n de loi $\mathcal{B}(n,p)$ qui représente le nombre de succès lors de n épreuves indépendantes.

3. Méthode utilisant le théorème central limite

Comme on utilise une approximation, cette méthode ne donne pas un véritable intervalle de confiance. Elle n'est valide que pour les grands n. On parle alors d'**intervalle de confiance**

asymptotique. On pose $Z = \frac{S_n - np}{\sqrt{npq}}$ qui suit approximativement la loi $\mathcal{N}(0,1)$. Pour faire le calcul, on va faire comme si la loi exacte était $\mathcal{N}(0,1)$.

Exemple numérique. On lance 1000 fois une pièce qui donne 523 fois pile. Trouver grâce au TCL au niveau de confiance 0.95 un intervalle de confiance pour $p = \mathbb{P}($ la pièce donne pile $)$.
On commence par chercher un réel $a > 0$ tel que $\mathbb{P}(|Z| \leq a) \approx 0.95$. Par symétrie de la courbe en cloche, on devrait avoir $\mathbb{P}(0 \leq Z \leq a) \approx \frac{0.95}{2} = 0.475$. D'après la table $\Psi(1.960) \approx 0.4750$. On choisit $a = 1.96$. On affirme donc que dans au moins 95% des cas $|Z| \leq 1,96$.
Ici la valeur expérimentale $S_{1000}(\omega)$ de la variable aléatoire de loi $\mathcal{B}(1000, p)$ est 523. Alors en substituant cette valeur à la place de S_n, dans au moins 95% des cas $|\frac{523-1000p}{\sqrt{1000pq}}| \leq 1,96$, c'est à dire $|523 - 1000p| \leq 1,96 \times \sqrt{1000pq}$.

Après division par 1000, $|0,523 - p| \leq 1,96 \times \frac{1}{\sqrt{1000}}\sqrt{pq}$.

On utilise maintenant l'astuce : pour tout p dans $[0,1]$, $p(1-p)$ est plus petit que $\frac{1}{4}$. (Pour le vérifier, on étudie les variations de la fonction trinôme $p \to p - p^2$ sur $[0,1]$ qui atteint son maximum au point $\frac{1}{2}$).

D'où $|0,523 - p| \leq 1,96 \times \frac{1}{\sqrt{1000}}\sqrt{\frac{1}{4}} \approx 0.01582$.
En utilisant le paragraphe fondamental de l'introduction, on voit que p appartient à l'intervalle de centre 0,523 de rayon 0.01582. On en déduit l'intervalle de confiance $I = [0.500.54]$ (noter que la borne inférieure a été arrondie par défaut pour rendre l'intervalle plus sûr, de même la borne supérieure a été arrondie par excès).

4. Méthode utilisant l'inégalité de Bienaymé-Tchebychev

Reprenons l'exemple numérique précédent en utilisant ce coup ci l'inégalité de Bienaymé-Tchebychev. À la différence du paragraphe précédent, on ne procède pas à une approximation donc cette méthode marche pour tous les entiers n, mais elle peut donner des intervalles peu intéressants.

Grâce à l'astuce $p(1-p) \leq \frac{1}{4}$, on déduit du corollaire 1 de l'inégalité de Bienaymé-Tchebychev que, $\mathbb{P}(|\frac{S_n}{n} - p)| > \beta) \leq \frac{1}{4n\beta^2}$.

Rappelons que l'inégalité de Bienaymé-Tchebychev sert à majorer la probabilité d'un événement anormal (qui va correspondre au complémentaire de l'intervalle de confiance).
Donc pour avoir une probabilité plus petite que 0.05 il suffit d'avoir $\frac{1}{4\times 1000\beta^2} \leq 0.05$, c'est à dire, $\beta \geq r := \sqrt{\frac{1}{200}} \approx 0,071$.

D'autre part, en reportant la valeur expérimentale $S_n(\omega) = 523$ dans la formule, on peut affirmer qu'avec une probabilité plus petite que 0.05, $|0,523 - p|$ est plus grand que β. En choisissant le meilleur β qui est $r \approx 0,071$ et en passant au complémentaire, on obtient l'intervalle de confiance pour p de centre 0,523 et de rayon 0,071.
On obtient l'intervalle $J = [0.452, 0.594]$ **plus mauvais** que I (J contient I).

Conclusion. Remarquer que les deux méthodes fournissent des intervalles de rayon proportionnel à $\frac{1}{\sqrt{n}}$, qui sont de plus en plus précis quand n croît. Toutefois, comme $\frac{1}{\sqrt{n}}$ décroît

assez lentement, ces intervalles ne sont pas extraordinaires. Pour diviser le rayon par 10, il faut centupler le nombre d'observations.